REPORT DOCUMENTATION PAGE

FORM APPROVED
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
02-25-97

3. REPORT TYPE AND DATES COVERED
FINAL 1-1-94 to 9-30-96

4. TITLE AND SUBTITLE OF REPORT
Rapid Solutions to Hard Problems Using Fast Messy Genetic Algorithms

5. FUNDING NUMBERS
G F 49620-94-0103

AFOSR-TR-97
0652

6. AUTHOR(S)
David E. Goldberg

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Illinois
506 South Wright Street
Urbana, IL 61801

8. PERFORMING ORGANIZATION REPORT NUMBER:
X0003

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 Duncan Avenue, Suite B115
Bolling AFB, DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER:
NK

11. SUPPLEMENTARY NOTES:
None

12a. DISTRIBUTION AVAILABILITY STATEMENT
None
Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT (Maximum 200 words)

This project developed and applied a type of non-traditional genetic algorithm called a fast messy genetic algorithm (fmGA). Critical bounding theory and computational experiements show that fmGAs converge to high quality solutions with high probability in times that grow no faster than a subquadratic function of the number of decision variables. These results have important ramifications for the design and operation of the next generation of Air Force systems.

14. SUBJECT TERMS
genetic algorithms, messy genetic algorithms, evolutionary computation

15. NUMBER OF PAGES:
8

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT:
unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
unclassified

20. LIMITATION OF ABSTRACT
UL

Final Report

Rapid Solutions to Hard Problems Using Fast Messy Genetic Algorithms AFOSR Grant No. F49620-94-1-0103

David E. Goldberg
Department of General Engineering
IlliGAL Report No. X0003 (Limited Circulation)
September 1997

Department of General Engineering
University of Illinois at Urbana-Champaign
117 Transportation Building
104 South Mathews Avenue
Urbana, IL 61801

Final Report for

Rapid Solutions to Hard Problems Using Fast Messy Genetic Algorithms, AFOSR Grant No. F49620-94-1-0103

1 January 1994 to 30 September 1996

David E. Goldberg
Department of General Engineering
The University of Illinois at Urbana-Champaign
117 Transportation Building
104 S. Mathews Avenue
Urbana, IL 61801

1 Introduction

This project (F49620-94-1-0103) has come to a close, and the record shows important advances as initially envisioned and beyond. Beside developing the fast messy GA and improving its thresholding mechanisms with theory and implementation improvements, two other linkage-friendly GAs have been demonstrated: the *gene expression messy genetic algorithm* (gemGA) (Kargupta, 1996) and the *linkage-learning genetic algorithm* (LLGA) (Harik & Goldberg, 1996). These latter efforts were initiated to make our methods simpler to implement and operate. The resulting codes use the principles of earlier fmGA work, but the mechanisms are considerably simpler and more robust.

Specifically, both Harik's LLGA and Kargupta's gemGA are solving difficult problems with 100s or 1000s of bits in linear to subquadratic times, although the two approaches appear to have complementary strengths. Kargupta's gemGA is fairly rapid in speculating on which building block is which, but it can be slowed when that initial judgment is in error. This can be disadvantageous when a problem is noisy, poorly scaled, or when the building blocks have considerable crosstalk. The linkage-learning GA is more deliberate in aggregating its building blocks. This can slow the algorithm when the problem is nearly separable, but can result in better solutions when noise, crosstalk, or bad scaling might fool the gemGA. Follow-on research has been funded to make a hybrid of the gemGA and LLGA to combine their strengths.

2 Major Accomplishments of the Project

This section highlights the key accomplishments of the project.

Fast messy GA study completed. Hillol Kargupta's dissertation demonstrating the efficiency and reliability of the fast messy GA in problems of bounded difficulty, including building block isolation, misleadingness, cross-talk, and adverse scaling, gives us confidence that these methods are ready for real-world application. His work applies the messy GA to the solution of a strategic defense radar tracking problem.

gemGA developed. Kargupta's study also developed a reorganized messy GA—a gene expression mGA or gemGA—that requires less manipulation of the thresholding and deletion schedules. In some cases, Kargupta is able to obtain linear time solutions to very difficult problems without problem specific knowledge.

General blackbox optimization framework developed. Kargupta has developed a new framework for considering similarities between all sampling blackbox optimizers (GA, SA, tabu search). The framework, called SEARCH (Search Envisioned As Relation and Class Hierachizing), decomposes the different functions of the population and then analyzes the limits of sampling time complexity. The framework generalizes the intuitions embodied in the fmGA and shows how other blackbox methods can be compared on a fair basis.

Linkage learning GA proved feasible. Georges Harik's dissertation has shown that direct recombinative learning of linkage in simple GAs is possible when crossover operators are combined with an mGA-like expresson mechanism. The dissertation is nearing completion and demonstrates fast solving in badly scaled problems. This is complementary to Kargupta's work, which appears to work best for BBs of uniform scale.

New population sizing opens several doors. A new, more accurate, model relating solution quality to GA population sizing has been developed using the gambler's ruin problem. This problem has been immediately useful for tightening complexity estimates of fmGA, gemGA, and LLGA complexity. It also is leading to first-time models of parallel GAs.

Convergence time noise delay determined. Selection intensity models have been constructed and solved for various selection schemes to calculate delay in convergence caused by noisy fitness functions. This result is useful in complexity estimates and for determining optimal sample sizes.

Optimal sample size determined. Using the noise delay results discussed above, the first integrated calculation of optimal sample size has been performed for GAs with noisy function evaluations. In noisy functions, a tradeoff often exists between taking a large number of samples and getting an accurate mean fitness value or taking a small number of samples and getting a less accurate value.

Coevolved niching invented. Two populations (businessmen and customers) are co-evolved to create clusters of solutions at diverse points throughout the search space. This technique differs from earlier versions of sharing in that the niche coverage is determined adaptively.

Integrated model of drift developed. New models of convergence in badly scaled problems have led to integrated models of drift. Theoretical computations and experiments agree well. This work is useful for showing why the LLGA is necessary—why simple GAs are poor at handling badly scaled problems—and for designing mechanisms to overcoming such difficulty.

Hybridization model constructed. A simple model of decision making in a hybrid scheme combining a selectorecombinative GA with a hillclimber has been constructed. Good agreement between theory and practice is obtained. This first effort will lead to a larger framework for understanding issues in building hybrids of GAs and other schemes.

3 Publications

The following is a list of publications submitted, accepted, and published with support from this project. In addition, MS and PhD theses supported by this project are listed at the end:

3.1 Submitted

Cantú-Paz, E., & Goldberg, D. E. (1997). *Modeling idealized bounding cases of parallel genetic algorithms.* Manuscript submitted for publication (Genetic Programming 97).

Cantú-Paz, E., & Goldberg, D. E. (1997). *Modeling speedups of idealized bounding cases of parallel genetic algorithms.* Manuscript submitted for publication (ICGA 97).

Harik, G., & Goldberg, D. E. (1996). *Learning linkage.* Presented at Foundations of Genetic Algorithms IV, University of San Diego, August 3–5, 1996. Under review for publication in refereed post-conference proceedings.

Horn, J. (1996). *Genetic algorithms (with sharing) in search, optimization, and machine learning.* Presented at Foundations of Genetic Algorithms IV, University of San Diego, August 3–5, 1996. Under review for publication in refereed post-conference proceedings.

Kargupta, H. (1996). *Constrained blackbox optimization: The SEARCH perspective.* Manuscript submitted for publication.

Kargupta, H. (1996). *Messy genetic algorithms: Recent developments.* Manuscript submitted for publication.

Kargupta, H., & Goldberg, D. E. (1996). *SEARCH: An alternative perspective toward blackbox optimization.* Manuscript submitted for publication (Journal of the ACM).

Kargupta, H., & Goldberg, D. E. (1996). *Blackbox optimization: Implications of SEARCH.* Manuscript submitted for publication (SIAM Journal of Computing).

Kargupta, H., & Goldberg, D. E. (1996). *SEARCH, blackbox optimization, and sample complexity.* Presented at Foundations of Genetic Algorithms IV, University of San Diego, August 3–5, 1996. Under review for publication in refereed post-conference proceedings.

Kargupta, H., Goldberg, D. E., & Wang, L. (1996). *Extending the class of order-k delineable problems for the gene expression messy genetic algorithm.* Manuscript submitted for publication.

Kargupta, H., Hanagandi, V., & Goldberg, D. E. (1996). *Unconstrained and constrained blackbox optimization: The SEARCH perspective.* Manuscript submitted to the Indian Journal of Computer Science.

3.2 Accepted but not published

Harik, G., Cantú-Paz, E., Goldberg, D. E., & Miller, B. L. (in press). The gambler's ruin problem, genetic algorithms, and the sizing of populations. *Proceedings of the 1997 IEEE International Conference on Evolutionary Computation.*

Weile, D. S., Michielssen, E., & Goldberg, D. E. (in press). Genetic algorithm design of Pareto optimal broad-band microwave absorbers. *Proceedings of the IEEE Symposium on Antennas and Propagation.*

Lobo, F., & Goldberg, D. E. (in press). Decision making in a hybrid genetic algorithm. *Proceedings of the 1997 IEEE International Conference on Evolutionary Computation.*

3.3 Published

Goldberg, D. E. (1995). The existential pleasures of genetic algorithms. In G. Winter, J. Périaux, M. Galán, & P. Cuesta (Eds.), *Genetic algorithms in engineering and computer science*, pp. 23–31. Chichester: John Wiley.

Goldberg, D. E. (1996). The design of innovating machines. *Computational Methods in Applied Sciences '96*, 100–104.

Goldberg, D. E. (1996). Toward a mechanics of conceptual machines. *Developments in Theoretical and Applied Mechanics XVIII*, 1–9.

Goldberg, D. E., & Harik, G. (1996). Case study in abnormal computational intelligence: Design of manufacturing and other anthropocentric systems. *International Journal of Computational Intelligence and Organizations, 1*(2), 78–93.

Harik, G. (1995). Finding multimodal solutions using restricted tournament selection. *Proceedings of the Sixth International Conference on Genetic Algorithms*, 24–31.

Horn, J., Goldberg, D. E., & Deb, K. (1994). Long path problems. *Parallel Problem Solving from Nature–PPSN III*, 149–158.

Horn, J, & Goldberg, D. E. (1995). Genetic algorithm difficulty and the modality of fitness landscapes. *Foundations of Genetic Algorithms 3*, 243–269.

Horn, J., & Goldberg, D. E. (1996). Natural niching for evolving cooperative classifiers. *Genetic Programming: Proceedings of the First Annual Conference*, 553–564.

Kargupta, H., Signal-to-noise, crosstalk, and long range problem difficulty in genetic algorithms. *Proceedings of the Sixth International Conference on Genetic Algorithms*, 193–200.

Kargupta, H. (1996). The gene expression messy genetic algorithm. *Proceedings of the 1996 IEEE Conference on Evolutionary Computation*, 814–819.

Kargupta, H. (1996). The performance of the gene expression messy genetic algorithm on real test functions. *Proceedings of the 1996 IEEE Conference on Evolutionary Computation*, 631–636.

Kargupta, H. (1996). The gene expression messy genetic algorithm for financial applications. *Proceedings of the IEEE/IAFE Conference on Computational Intelligence for Financial Engineering*, 155-161.

Kargupta, H. (1996). SEARCH and a computational perspective of evolution. *Proceedings of Artificial Life V*, 56–63.

Kargupta, H., & Goldberg D. E. (1996). Polynomial complexity blackbox search: Lessons from the SEARCH framework. *Proceedings of the 1996 IEEE Conference on Evolutionary Computation*, 792–797.

Mahfoud, S. W. (1995). A comparison of parallel and sequential niching methods. *Proceedings of the Sixth International Conference on Genetic Algorithms*, 136–143.

Mahfoud, S. W.& Goldberg, D. E. (1995). Parallel recombinative simulated annealing: A genetic algorithm. *Parallel Computing, 21*, 1–28.

Miller, B., & Goldberg, D. E. (1995). Genetic algorithms, tournament selection, and the effects of noise. *Complex Systems, 9*, 193–212.

Miller, B., & Goldberg, D. E. (1996). Genetic algorithms, selection schemes, and the varying effects of noise. *Evolutionary Computation 4*(2), 113-131.

Miller, B., & Goldberg, D. E. (1996). Optimal sampling for genetic algorithms. *Proceedings of Artificial Neural Networks in Engineering*, 1–7.

Miller, B. & Shaw, M. (1996). Genetic algorithms with dynamic niche sharing for multimodal function optimization. *Proceedings of the 1996 IEEE Conference on Evolutionary Computation.*

Thierens, D., & Goldberg, D. E. (1994). Convergence models of genetic algorithm selection schemes. *Parallel Problem Solving from Nature–PPSN III*, 117–129.

Weile, D. S. Michielssen, E., & Goldberg, D. E. (1996). Pareto optimal design of microwave absorbers. *IEEE Transactions on Electromagetic Compatability, 38*(3), 518–525.

3.4 M.S. Theses Completed with Air Force Support

Horn, J. (1995). *Genetic algorithms, problem difficulty, and the modality of fitness landscapes* (IlliGAL Report No. 95004 and M.S. thesis, Computer Science). Urbana: University of Illinois at Urbana-Champaign.

Wilcox, J. (1995). *Organizational learning within a learning classifier system* (IlliGAL Report No. 95003 and M.S. thesis, Computer Science). Urbana: University of Illinois at Urbana-Champaign.

3.5 Ph.D. Dissertations

In addition to the dissertations listed below, dissertations by G. Harik and B. Miller are expected to be completed by May. These dissertations were in part support by this grant.

Kargupta, H. (1995). *SEARCH, polynomial complexity, and the fast messy genetic algorithm* (IlliGAL Report No. 95008 and PhD dissertation, Computer Science). Urbana: University of Illinois at Urbana-Champaign.

Mahfoud, S. (1995). *Niching methods for genetic algorithms* (IlliGAL Report No. 95001 and Ph.D. dissertation, Computer Science). Urbana: University of Illinois at Urbana-Champaign.

Thierens, D. (1995). *Analysis and design of genetic algorithms* (Ph.D. dissertation). Leuven: Katholieke Universiteit Leuven.

3.6 Transitions

Electromagnetic applications. I have worked with E. Michielssen here at the University of Illinois on applying multiobjective GAs to the solution of antenna and absorber design. Professor Michielssen is under contract to AFOSR with a project "Electromagnetic Scattering from Complex Structures."

Laser-hardened materials. Dr. Ruth Pachter at Wright-Patterson Air Force Base would like to design laser-resistant materials to protect pilot vision during combat. To do so requires, an accurate means of predicting secondary structure of complex molecules. Dr. Pachter has worked with simple and messy GAs with some success. I met with her on 23 August 1995 at WPAFB to discuss her application. This contact is continuing as part of the new grant.

GA basic and applied research. A continuing interchange with Dr. Gary Lamont at the Air Force Institute of Technology has resulted in a corpus of new basic and applied GA work by a growing number of Air Force officers with advanced degrees. Captain Merkle is the latest product of thise effort, finishing an important dissertation analyzing the primordial phase of the fast messy GA; however, work is continuing on a number of OR, scheduling, neural net, and other projects.

John Deere, manufacturing, and people in the loop. As part of a GA project for John Deere, I constructed some models of human behavior that help explain the need for just-in-time or lean methods of manufacturing without resorting to excessive penalties on inventory. Deere is considering the ramifications of these models for manufacturing system design.

GA stock trading. Sam Mahfoud has taken his Air Force supported training and is now developing GA code in a stock portfolio management system at LBS Capital Management in Clearwater, Florida. A number of the techniques developed under this work are finding their way into these new systems.

Data mining and financial applications. Hillol Kargupta has applied techniques derived from this project to financial systems and data mining applications while at Los Alamos National Laboratory as part of his postdoc there.

Wired, WSJ, Newsweek, GASOML, and indirect transitions. In the past several years, GA applications have exploded. My work has played an important role in both popularizing GAs (especially my earlier *Genetic Algorithms in Search, Optimization, and Machine Learning* and legitimizing the field (especially my foundational work on applicable theory and messy GAs sponsored by AFOSR). It is difficult to measure this more diffuse kind of transition than one in which a person plays a direct role; however, the recent pubication of articles in *Newsweek*, the *Wall Street Journal*, and *Wired* (February, 1997) are important signposts of a form of indirect intellectual leverage or transition that this AFOSR sponsored work is having.

DARPA Image Recognition System. A practical system for image recognition has been enhanced by applying a messy GA. Darrell Whitley at Colorado State University took mGA technology out of our literature and applied it to an established state-of-the-art image recognition system in a detector formation subsystem. The messy GA, outperformed previous best-performance algorithms by a wide margin. This system is part of a large autonomous vehicle project and the performance boost will have practical consequences.

Evolvable Hardware. A team led by Tetsuya Higuchi at MITI's Electrotechnical Laboratory (ETL) in Tsukuba, Japan has used a messy GA (renamed a VGA or variable GA) to program field programmable gate arrays (FPGAs) and programmable logic devices (PLDs) for a number of pattern recognition and fault-tolerant system applications. Evolvable hardware is a growing application area for GAs.